

6

PATENT ABSTRACTS OF JAPAN

(11) Publication number : 2001-331398
 (43) Date of publication of application : 30.11.2001

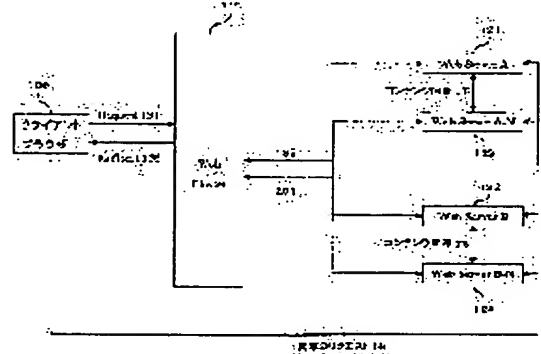
(51) Int. Cl. G06F 13/00
 G06F 9/46
 G06F 12/00

(21) Application number : 2000-148710 (71) Applicant : OPRO JAPAN CO LTD
 (22) Date of filing : 19.05.2000 (72) Inventor : SATOMI KAZUNORI
 KOIKE YASUTSUGU

(54) SERVER-MANAGING SYSTEM

(57) Abstract:

PROBLEM TO BE SOLVED: To provide a server-managing system capable of performing efficient distribution processing.
 SOLUTION: A web cluster 110 can access web servers 121-124 and inquires system load information at prescribed time intervals. When an access request 131 is received from a client browser 100, the web server of the lowest system load information is selected, and redirect information 133 describing the identifier of the relevant web server is returned to the client browser 100. The client browser 100, which receives the redirect information 133, sends a real request 141 to the relevant web server, on the basis of the identifier described on that information 133.



LEGAL STATUS

[Date of request for examination] 19.05.2000
 [Date of sending the examiner's decision of rejection]
 [Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]
 [Date of final disposal for application]
 [Patent number]
 [Date of registration]
 [Number of appeal against examiner's decision of rejection]
 [Date of requesting appeal against examiner's decision of rejection]
 [Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2001-331398

(P2001-331398A)

(43) 公開日 平成13年11月30日 (2001. 11. 30)

(51) Int.Cl. ⁷	識別記号	F I	テ-マコード [*] (参考)
G 0 6 F 13/00	3 5 7	G 0 6 F 13/00	3 5 7 Z 5 B 0 8 2
9/46	3 6 0	9/46	3 6 0 C 5 B 0 8 9
12/00	5 4 6	12/00	5 4 6 A 5 B 0 9 8

審査請求 有 請求項の数13 O L (全 8 頁)

(21) 出願番号 特願2000-148710 (P2000-148710)

(22) 出願日 平成12年 5 月19日 (2000. 5. 19)

(71) 出願人 397053292

日本オプロ株式会社

東京都品川区西五反田 8-8-15

(72) 発明者 里見 一典

東京都品川区西五反田 8-8-15 日本オ
プロ株式会社内

(72) 発明者 小池 康嗣

東京都品川区西五反田 8-8-15 日本オ
プロ株式会社内

(74) 代理人 100086368

弁理士 萩原 誠

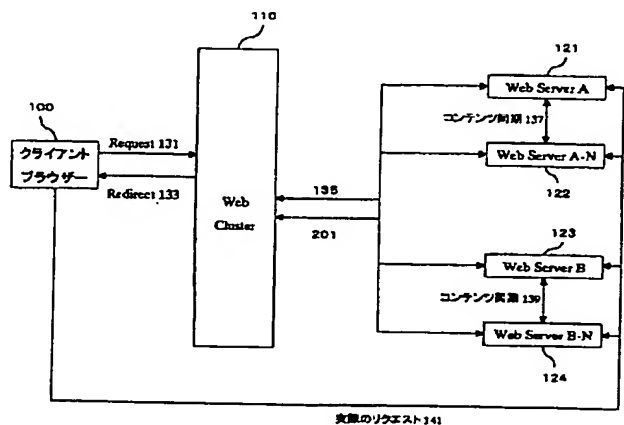
最終頁に続く

(54) 【発明の名称】 サーバ管理システム

(57) 【要約】

【課題】 効率の良い分散処理を行うことのできるサーバ管理システムを提供すること。

【解決手段】 ウェブクラスタ 110 は、ウェブサーバ 121~124 とアクセスすることが可能であり、所定の時間間隔でシステム負荷情報を問合せている。そして、クライアントブラウザ 100 からアクセス要求 131 を受け取った際、システム負荷情報の最も低いウェブサーバを選択し、当該ウェブサーバの識別子を載せたリダイレクト情報 133 をクライアントブラウザ 100 に返信する。リダイレクト情報 133 を受け取ったクライアントブラウザ 100 は、その情報 133 に載せられてきた識別子に基づいて当該ウェブサーバへ実際のリクエスト 141 を送出する。



【特許請求の範囲】

【請求項1】 複数のウェブサーバのシステム負荷分散、及びそれらのコンテンツの整合性をとるサーバ管理システムであり、

コンピュータネットワークで接続されたクライアント及びウェブサーバと、前記ウェブサーバと前記クライアントとの間に介在するウェブサーバ管理装置とで構成され、

前記ウェブサーバ管理装置は、

前記複数のウェブサーバとアクセスするアクセス手段と、

所定の時間間隔で各ウェブサーバにシステム負荷情報を問合わせる手段と、

各ウェブサーバのシステム負荷情報を記憶する記憶手段と、

前記クライアントからアクセス要求を受け取った際、前記記憶手段を参照し、システム負荷情報の最も低いウェブサーバを選択する選択手段と、

選択一致したウェブサーバの識別子を載せたリダイレクト情報を前記クライアントに返信する手段とを具備し、前記複数のウェブサーバは、

前記ウェブサーバ管理装置からシステム負荷情報の問合わせを受けた際、

自身のシステム負荷情報を前記ウェブサーバ管理装置に送出する手段を具備することを特徴とするサーバ管理システム。

【請求項2】 請求項1に記載のサーバ管理システムにおいて、

前記システム負荷情報の問合せに対する応答が所定時間内にない場合、または、前記問合せに対してシステムダウンを意味する応答が返ってきた場合、

前記ウェブサーバ管理装置は、

当該ウェブサーバがシステムダウンしているものとみなし、

当該ウェブサーバが復旧したことを確認するまで、そのシステム負荷情報を前記記憶手段から削除することを特徴とするサーバ管理システム。

【請求項3】 請求項1または請求項2に記載のサーバ管理システムにおいて、

前記ウェブサーバ管理装置は、

アクセス数の多いコンテンツのうち、時間的変化の比較的少ないコンテンツを一時的に保持する記憶領域を有し、

ウェブサーバから前記コンテンツの更新情報を受信するまで、

ウェブサーバに代わってクライアントに前記コンテンツを送信することを特徴とするサーバ管理システム。

【請求項4】 請求項1乃至請求項3のいずれかに記載のサーバ管理システムにおいて、

前記複数のウェブサーバ間で整合性をとるように指定さ

れた整合指定コンテンツがある場合、

各ウェブサーバの整合指定コンテンツを監視する監視手段を具備し、

いずれかのウェブサーバの整合指定コンテンツに変更が確認された場合、

変更された前記整合指定コンテンツを、他のウェブサーバに自動コピーすることを特徴とするサーバ管理システム。

【請求項5】 コンピュータネットワークで接続されたクライアント及びウェブサーバと、前記ウェブサーバと前記クライアントとの間に介在するウェブサーバ管理装置とを用いて、複数のウェブサーバのシステム負荷分散、及びそれらのコンテンツの整合性をとるサーバ管理方法であり、

前記ウェブサーバ管理装置は、

前記複数のウェブサーバとアクセスし、

所定の時間間隔で各ウェブサーバにシステム負荷情報を問合わせ、

各ウェブサーバのシステム負荷情報を記憶領域に記憶し、

前記クライアントからアクセス要求を受け取った際、

前記記憶領域を参照し、システム負荷情報の最も低いウェブサーバを選択し、

選択一致したウェブサーバの識別子を載せたリダイレクト情報を前記クライアントに返信し、

前記複数のウェブサーバは、

前記ウェブサーバ管理装置からシステム負荷情報の問合わせを受けた際、

自身のシステム負荷情報を前記ウェブサーバ管理装置に送出することを特徴とするサーバ管理方法。

【請求項6】 請求項5に記載のサーバ管理方法において、

前記システム負荷情報の問合せに対する応答が所定時間内にない場合、または、前記問合せに対してシステムダウンを意味する応答が返ってきた場合、

前記ウェブサーバ管理装置は、

当該ウェブサーバがシステムダウンしているものとみなし、

当該ウェブサーバが復旧したことを確認するまで、そのシステム負荷情報を前記記憶領域から削除することを特徴とするサーバ管理方法。

【請求項7】 請求項5または請求項6に記載のサーバ管理方法において、前記ウェブサーバ管理装置は、アクセス数の多いコンテンツのうち、時間的変化の比較的少ないコンテンツを一時的に保持し、

ウェブサーバから前記コンテンツの更新情報を受信するまで、

ウェブサーバに代わってクライアントに前記コンテンツを送信することを特徴とするサーバ管理方法。

【請求項8】 請求項5乃至請求項7のいずれかに記載

のサーバ管理方法において、
前記複数のウェブサーバ間で整合性をとるように指定された整合指定コンテンツがある場合、

各ウェブサーバの整合指定コンテンツを監視し、
いずれかのウェブサーバの整合指定コンテンツに変更が確認された場合、

変更された前記整合指定コンテンツを、他のウェブサーバに自動コピーすることを特徴とするサーバ管理方法。

【請求項 9】 複数のウェブサーバとアクセスする第 1 の手順と、

所定の時間間隔で各ウェブサーバにシステム負荷情報を問合わせる第 2 の手順と、

各ウェブサーバのシステム負荷情報を記憶領域に記憶する第 3 の手順と、

クライアントからアクセス要求を受け取った際、

前記記憶領域を参照し、システム負荷情報の最も低いウェブサーバを選択する第 4 の手順と、

選択一致したウェブサーバの識別子を載せたリダイレクト情報を前記クライアントに返信する第 5 の手順とを実行するソフトウェアを蓄積したことを特徴とする電子媒体。

【請求項 10】 請求項 9 に記載の電子媒体において、前記システム負荷情報の問合せに対する応答が所定時間内にない場合、または、前記問合せに対してシステムダウンを意味する応答が返ってきた場合、当該ウェブサーバがシステムダウンしているものとみなし、

当該ウェブサーバが復旧したことを確認するまで、そのシステム負荷情報を前記記憶領域から削除する手順を実行するソフトウェアを蓄積したことを特徴とする電子媒体。

【請求項 11】 請求項 9 または請求項 10 に記載の電子媒体において、

アクセス数の多いコンテンツのうち、時間的変化の比較的少ないコンテンツを一時的に保持する第 1 の手順と、ウェブサーバから前記コンテンツの更新情報を受信するまで、

ウェブサーバに代わってクライアントに前記コンテンツを送信する第 2 の手順とを実行するソフトウェアを蓄積したことを特徴とする電子媒体。

【請求項 12】 請求項 9 乃至請求項 11 のいずれかに記載の電子媒体において、

前記複数のウェブサーバ間で整合性をとるように指定された整合指定コンテンツがある場合、

各ウェブサーバの整合指定コンテンツを監視する第 1 の手順と、

いずれかのウェブサーバの整合指定コンテンツに変更が確認された場合、

変更された前記整合指定コンテンツを、他のウェブサーバに自動コピーする第 2 の手順とを実行するソフトウェ

アを蓄積したことを特徴とする電子媒体。

【請求項 13】 システム負荷情報の問合わせを受けた際、

自身のシステム負荷情報を送出する手順を実行するソフトウェアを蓄積したことを特徴とする電子媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明はサーバ管理システムに関するものであり、より具体的には、複数のウェブサーバのシステム負荷分散、及びそれらのコンテンツの整合性をとるサーバ管理システムに関する。

【0002】

【従来の技術】WWWシステムにおけるクライアントサーバシステムは、サーバ側であるウェブサーバとクライアント側であるクライアントブラウザとで構成されており、ウェブサーバはクライアントブラウザからの要求に応じて、HTML 文書や画像などの情報をインターネットなどのコンピュータネットワークを通じて送信している。そして、情報提供者であるウェブサーバの管理者にとって、クライアントブラウザからのアクセス数が増加することは一般的に好ましいことである。

【0003】

【発明が解決しようとする課題】しかしながら、同時に多くのアクセスがあった場合、当該ウェブサーバの処理速度が極端に低下してしまうという事態が生じる。このような事態を回避するため、ネットワーク上の資源を有効に活用し、効率良くアプリケーションを実行することは重要な課題であり、そのためにはネットワーク上の資源の状態を正しく把握する必要がある。本発明は、上記課題を解決するため、複数あるウェブサーバの負荷情報を把握し、システム負荷を分散することにより効率の良い分散処理を行うことのできるサーバ管理システムを提供することを目的とする。

【0004】

【課題を解決するための手段】上記課題を解決するため、本発明は、コンピュータネットワークで接続されたクライアント及びウェブサーバと、ウェブサーバとクライアントとの間に介在するウェブサーバ管理装置とで構成され、ウェブサーバ管理装置は、複数のウェブサーバとアクセスするアクセス手段と、所定の時間間隔で各ウェブサーバにシステム負荷情報を問合わせる手段と、各ウェブサーバのシステム負荷情報を記憶する記憶手段と、クライアントからアクセス要求を受け取った際、記憶手段を参照し、システム負荷情報の最も低いウェブサーバを選択する選択手段と、選択一致したウェブサーバの識別子を載せたリダイレクト情報をクライアントに返信する手段とを具備し、複数のウェブサーバは、ウェブサーバ管理装置からシステム負荷情報の問合わせを受けた際、自身のシステム負荷情報をウェブサーバ管理装置に送出する手段を具備することを特徴とするサーバ管理

システムを提供する。

【0005】また、上記サーバ管理システムにおいて、システム負荷情報の問合せに対する応答が所定時間内にない場合、または、問合せに対してシステムダウンを意味する応答が返ってきた場合、ウェブサーバ管理装置は、当該ウェブサーバがシステムダウンしているものとみなし、当該ウェブサーバが復旧したことを確認するまで、そのシステム負荷情報を記憶手段から削除することを特徴とする。また、上記サーバ管理システムにおいて、ウェブサーバ管理装置は、アクセス数の多いコンテンツのうち、時間的変化の比較的小さいコンテンツを一時的に保持する記憶領域を有し、ウェブサーバからコンテンツの更新情報を受信するまで、ウェブサーバに代わってクライアントにコンテンツを送信することを特徴とする。また、上記サーバ管理システムは、複数のウェブサーバ間で整合性をとるように指定された整合指定コンテンツがある場合、各ウェブサーバの整合指定コンテンツを監視する監視手段を具備し、いずれかのウェブサーバの整合指定コンテンツに変更が確認された場合、変更された整合指定コンテンツを、他のウェブサーバに自動コピーすることを特徴とする。

【0006】さらに、本発明は、コンピュータネットワークで接続されたクライアント及びウェブサーバと、ウェブサーバとクライアントとの間に介在するウェブサーバ管理装置とを用いて、複数のウェブサーバのシステム負荷分散、及びそれらのコンテンツの整合性をとるサーバ管理方法を提供し、ウェブサーバ管理装置は、複数のウェブサーバとアクセスし、所定の時間間隔で各ウェブサーバにシステム負荷情報を問い合わせ、各ウェブサーバのシステム負荷情報を記憶領域に記憶し、クライアントからアクセス要求を受け取った際、記憶領域を参照し、システム負荷情報の最も低いウェブサーバを選択し、選択一致したウェブサーバの識別子を載せたリダイレクト情報をクライアントに返信し、複数のウェブサーバは、ウェブサーバ管理装置からシステム負荷情報の問い合わせを受けた際、自身のシステム負荷情報をウェブサーバ管理装置に送出することを特徴とするサーバ管理方法を提供する。

【0007】また、上記サーバ管理方法において、システム負荷情報の問合せに対する応答が所定時間内にない場合、または、問合せに対してシステムダウンを意味する応答が返ってきた場合、ウェブサーバ管理装置は、当該ウェブサーバがシステムダウンしているものとみなし、当該ウェブサーバが復旧したことを確認するまで、そのシステム負荷情報を記憶領域から削除することを特徴とする。また、上記サーバ管理方法において、ウェブサーバ管理装置は、アクセス数の多いコンテンツのうち、時間的変化の比較的小さいコンテンツを一時的に保持し、ウェブサーバからコンテンツの更新情報を受信するまで、ウェブサーバに代わってクライアントにコンテ

ンツを送信することを特徴とする。また、上記サーバ管理方法は、複数のウェブサーバ間で整合性をとるように指定された整合指定コンテンツがある場合、各ウェブサーバの整合指定コンテンツを監視し、いずれかのウェブサーバの整合指定コンテンツに変更が確認された場合、変更された整合指定コンテンツを、他のウェブサーバに自動コピーすることを特徴とする。

【0008】さらに、本発明は、複数のウェブサーバとアクセスする第1の手順と、所定の時間間隔で各ウェブサーバにシステム負荷情報を問い合わせる第2の手順と、各ウェブサーバのシステム負荷情報を記憶領域に記憶する第3の手順と、クライアントからアクセス要求を受け取った際、記憶領域を参照し、システム負荷情報の最も低いウェブサーバを選択する第4の手順と、選択一致したウェブサーバの識別子を載せたリダイレクト情報をクライアントに返信する第5の手順とを実行するソフトウェアを蓄積したことを特徴とする電子媒体を提供する。

【0009】また、上記電子媒体は、システム負荷情報の問合せに対する応答が所定時間内にない場合、または、問合せに対してシステムダウンを意味する応答が返ってきた場合、当該ウェブサーバがシステムダウンしているものとみなし、当該ウェブサーバが復旧したことを確認するまで、そのシステム負荷情報を記憶領域から削除する手順を実行するソフトウェアを蓄積したことを特徴とする。また、上記電子媒体は、アクセス数の多いコンテンツのうち、時間的変化の比較的小さいコンテンツを一時的に保持する第1の手順と、ウェブサーバからコンテンツの更新情報を受信するまで、ウェブサーバに代わってクライアントにコンテンツを送信する第2の手順とを実行するソフトウェアを蓄積したことを特徴とする。

【0010】また、上記電子媒体は、複数のウェブサーバ間で整合性をとるように指定された整合指定コンテンツがある場合、各ウェブサーバの整合指定コンテンツを監視する第1の手順と、いずれかのウェブサーバの整合指定コンテンツに変更が確認された場合、変更された整合指定コンテンツを、他のウェブサーバに自動コピーする第2の手順とを実行するソフトウェアを蓄積したことを特徴とする。さらに、本発明は、システム負荷情報の問い合わせを受けた際、自身のシステム負荷情報を送出する手順を実行するソフトウェアを蓄積したことを特徴とする電子媒体を提供する。

【0011】

【発明の実施の形態】以下本発明の実施の形態について説明する。図1は本発明のサーバ管理システムを説明するシステム概略図、図2は本発明のサーバ管理システムの全体構成図である。また、図3及び図4は本発明のウェブサーバ管理装置内の処理フロー図、図5はコンテンツ監視プログラムを説明するための処理フロー図、図6は負荷情報のデータ構成図である。

【0012】図1において、本発明のサーバ管理システムは、クライアントブラウザ100、複数のウェブサーバ121～124（Web Server A～Web Server B-N）、及びウェブクラスタ（Web Cluster）110で構成されている。本説明では、ウェブサーバ管理装置をウェブクラスタ110と称す。ウェブクラスタ110は、ウェブサーバ121～124とアクセスすることが可能であり、所定の時間間隔でシステム負荷情報を問合せている。そして、クライアントブラウザ100からアクセス要求（Request）131を受け取った際、システム負荷情報の最も低いウェブサーバを選択し、当該ウェブサーバの識別子を載せたリダイレクト情報（Redirect）133をクライアントブラウザ100に返信する。リダイレクト情報133を受け取ったクライアントブラウザ100は、その情報133に載せられてきた識別子に基づいて当該ウェブサーバへ実際のリクエスト141を送出する。

【0013】一方、ウェブサーバ121～124は、ウェブクラスタ110からシステム負荷情報の問合わせを受けた際、自身のシステム負荷情報135をウェブクラスタ110に送出する。本実施の形態では、ウェブサーバ121及び122（以下、サーバA群と称する）はメインサブの関係になっており、整合性をとるように指定されたコンテンツAを各自保持しているとする。さらに、ウェブサーバ123及び124（以下、サーバB群と称する）もメインサブの関係になっており、整合性をとるように指定されたコンテンツBを各自保持しているとする。そのため、サーバA群間ではコンテンツ同期137がとられ、サーバB群間ではコンテンツ同期139がとられている。

【0014】次に、図2を用いてウェブクラスタ110とサーバA群とのデータのやり取りを詳しく説明する。サーバA群を構成するウェブサーバ121と122とは、整合性をとるように指定されたコンテンツA、負荷情報確認プログラム、及びコンテンツ監視プログラムとをそれぞれ有している。一方、ウェブクラスタ110はウェブサーバ121～124のシステム負荷情報を記憶した負荷情報203、及びキャッシュエリアコンテンツ情報205を有している。ここで、キャッシュエリアコンテンツ情報205とは、アクセス数の多いコンテンツのうち時間的変化の比較的小さいコンテンツを記憶した記憶領域のことである。負荷情報確認プログラムの各実行結果は、システム負荷情報135としてウェブクラスタ110の負荷情報203に記憶される。また、コンテンツ監視プログラムによって、キャッシュエリアコンテンツ情報205に相当するコンテンツの変更が確認されたときも、キャッシュ（Cache）登録・更新情報201がキャッシュエリアコンテンツ情報205に記憶される。

【0015】次に、図3を用いてウェブクラスタ110

がウェブサーバ121～124にシステム負荷情報の問合わせを行うときの処理を説明する。所定時間が経過すると、本処理は開始し（ステップ301）、ウェブクラスタ110はウェブサーバ121～124に対してシステム負荷情報の問合わせを行う（ステップ303）。ステップ307において応答がなければ、当該ウェブサーバのシステム負荷情報を負荷情報203から削除して（ステップ315）処理を終了する（ステップ317）。

【0016】一方、ステップ307において応答があれば、その応答がシステムダウンを意味するものであるか否か確認する（ステップ311）。確認の結果、システムダウンを意味していれば当該ウェブサーバのシステム負荷情報を負荷情報203から削除して（ステップ315）処理を終了する（ステップ317）。一方、正常に動いていることを意味する応答が返ってくれば、当該システム負荷情報を負荷情報203に記憶して（ステップ313）処理を終了する（ステップ317）。

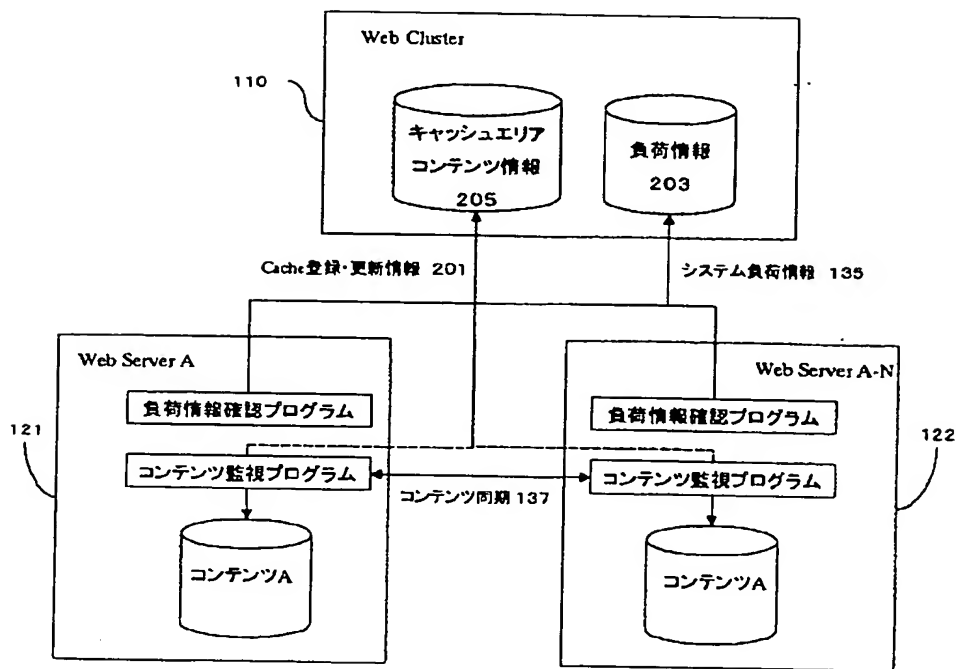
【0017】図3のような手順にしたがって取得した負荷情報203の一例を図6に示す。列610には121～124のウェブサーバ名、列620には各ウェブサーバのシステム負荷情報が記憶されている。よって、行601～604にはウェブサーバ名とそのウェブサーバのシステム負荷情報とが対応づけて記憶されている。この状態で、ウェブブラウザ100よりウェブクラスタ110にアクセス要求131が出されたときの処理フロー図を図4に示す。

【0018】図4において、本処理が開始すると（ステップ401）、ウェブクラスタ110は負荷情報203（ここでは図6）を参照する（ステップ403）。システム負荷情報の最小値は行602の「4」であり、対応するウェブサーバは122（Web Server A-N）である。よって、ウェブクラスタ110はウェブサーバ122を選択し（ステップ405）リダイレクト情報に載せる。そして、そのリダイレクト情報をクライアントブラウザ100に返信し（ステップ407）、処理を終了する（ステップ409）。

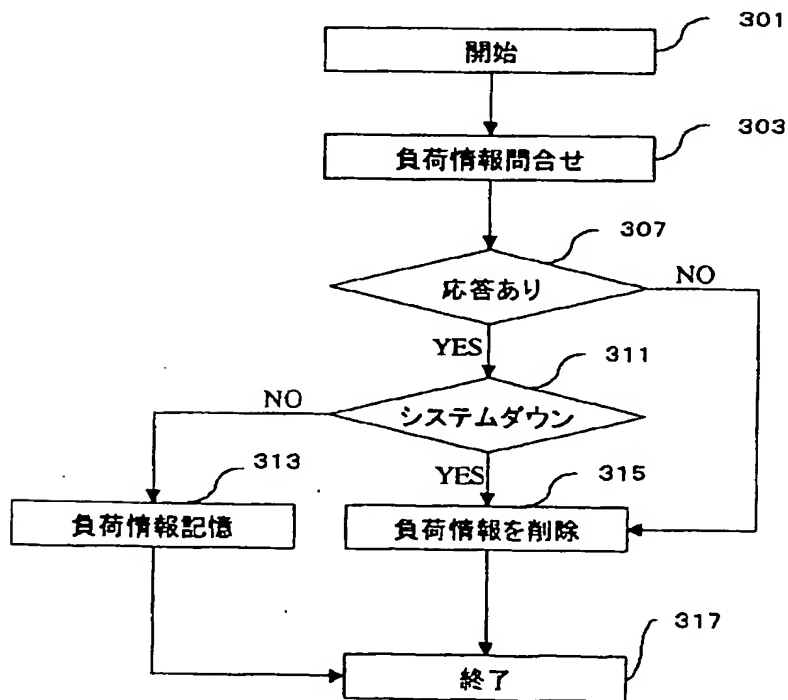
【0019】次に、図5を用いてコンテンツ監視プログラムの処理を説明する。本実施の形態のサーバA群間では、コンテンツAに整合性をもたせることが指定されている。そのため、コンテンツ監視プログラムが開始すると（ステップ501）、コンテンツAが変更されるか否か監視する（ステップ503）。ここで、変更が確認されない場合にはそのまま処理を終了する（ステップ515）。確認の結果、ウェブサーバ121のコンテンツAに変更が確認された場合には変更部分がキャッシュ（Cache）情報であるか否か確認する（ステップ505）。

【0020】変更部分がキャッシュ情報でない場合には、ウェブサーバ121のコンテンツAを、ウェブサーバ122のコンテンツAに自動コピーするなどして更新する（ステップ507）。一方、変更部分がキャッシュ

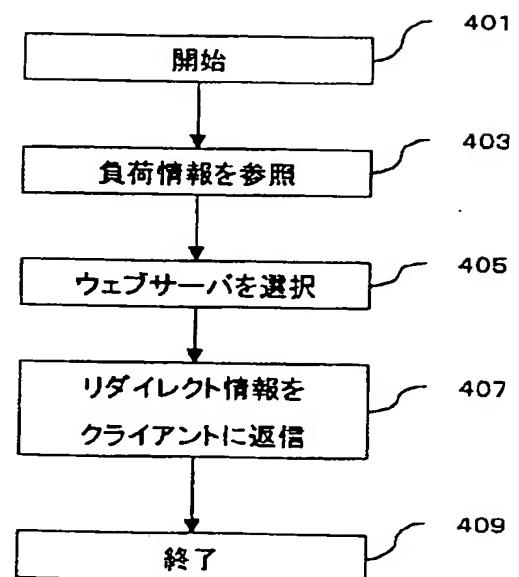
【図2】



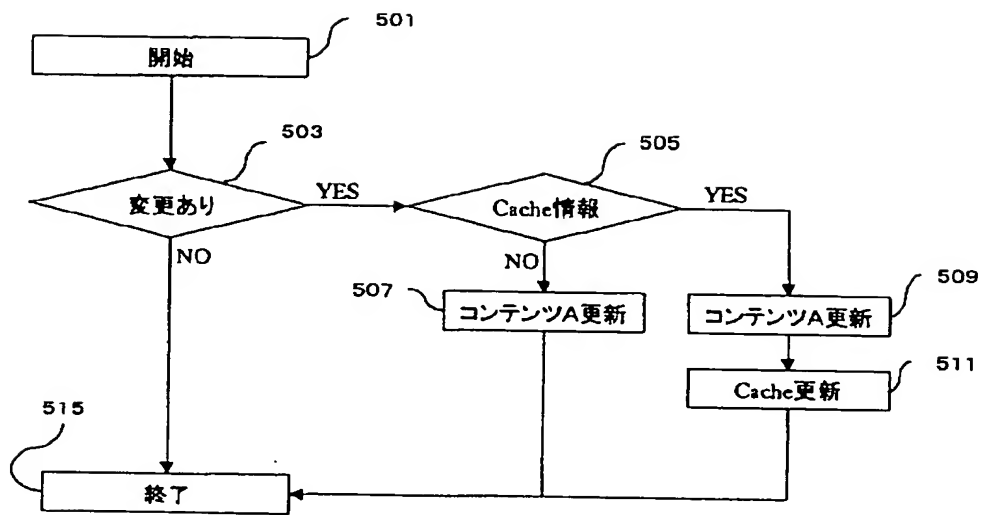
【図3】



【図4】



【図5】



【図6】

ウェブサーバ	システム負荷情報
Web Server A	8
Web Server A-N	4
Web Server B	6
Web Server B-N	5

フロントページの続き

Fターム(参考) 5B082 HA05 HA08
 5B089 GA11 GA19 GB02 HA10 JA21
 KA06 KA13 KC23 KC30 MA03
 5B098 AA10 GC10 GD02 GD14